TEORÍA DE LA PROBABILIDAD

Medida generada por una función de distribución

Introducción

Se ha extendido mucho la idea de que la teoría de la medida es algo muy complicado; sin embargo, como el mismo nombre lo indica, se trata de una teoría que nos brinda métodos para medir. En la teoría de la probabilidad se utiliza para encontrar probabilidades.

Recordemos que actualmente, para el estudio de algún fenómeno, el cual, ya sea por su complejidad o por alguna razón que se asume intrínseca, se modela pensando que algunos de los hechos que lo componen ocurren al azar. Para esto se introduce el concepto de experimento aleatorio, el cual tiene la característica de que su resultado no está únicamente determinado, sino que puede ser alguno de los elementos del conjunto de todos sus posibles resultados, el cual es llamado el espacio muestral del experimento y se le suele denotar por las letra Ω.

También se introduce el concepto de evento, el cual consiste de cualquier aseveración acerca del resultado del experimento antes de que éste se realice. Cada evento se representa como un subconjunto del espacio muestral, a saber, el formado por todos los posibles resultados que hacen que la aseveración resulte verdadera. Se trata entonces de calcular la probabilidad de ocurrencia de cada uno de los eventos; es decir, se busca encontrar una función P, definida sobre el conjunto de eventos y con valores entre 0 y 1, la cual asigna a cada evento su probabilidad de ocurrencia. De esta forma, cada experimento aleatorio se modela mediante una terna $(\Omega, \Im, P)$, donde $\Im$ es el conjunto formado por todos los eventos.

En el desarrollo histórico de la teoría de la probabilidad se eligió buscar una función de probabilidad que tenga las siguientes características:

i) $P(A) \geq 0$ para cualquier $A \in \Im$.

ii) $P(\Omega) = 1$.

iii) P es $\sigma-$aditiva (sigma aditiva); es decir, si $(A_n)_{n\in\mathbb{N}}$ es una sucesión de eventos tales que cualquier pareja de ellos está formada por conjuntos ajenos, entonces:

$$P\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

Una limitación que tiene este modelo es que no siempre es posible encontrar una función P, con las características anteriores, la cual esté definida sobre todos los subconjuntos del espacio muestral. Es por esto que se introduce el concepto de $\sigma-$álgebra (sigma álgebra), ya que, si bien P no está definida sobre todos los subconjuntos de Ω, el conjunto de eventos, es decir de los subconjuntos de Ω a los cuales sí es posible asignarle una probabilidad de ocurrencia, tiene las siguientes propiedades:

$i)$ $\Omega \in \Im$.

$ii)$ Si $A \in \Im$, entonces $A^c \in \Im$.

$iii)$ Si $(A_n)_{n\in\mathbb{N}}$ es una sucesión de elementos de $\Im$, entonces $\cup_{n=1}^{\infty} A_n \in \Im$.

Cuando se tienen estas propiedades se dice que $\Im$ es una $\sigma-$álgebra de subconjuntos de Ω.

A una terna $(\Omega, \Im, P)$ con las propiedades antes mencionadas se le llama un espacio de probabilidad y a la función P se le llama una medida de probabilidad.

En general, cuando se tiene una experimento aleatorio, el primer problema a resolver consiste en encontrar un espacio de probabilidad que lo represente.

Una vez que se tiene definido un espacio de probabilidad $(\Omega, \Im, P)$, se introduce el concepto de variable aleatoria real, la cual se define como cualquier función $X : \Omega \to \mathbb{R}$ con la propiedad de que, para cualquier $x \in \mathbb{R}$, el conjunto de elementos $\omega \in \Omega$ tales que $X(\omega) \leq x$, lo cual se denota por $[X \leq x]$, sea un evento, es decir, sea un elemento de $\Im$.

Dada una variable aleatoria X se define su función de distribución $F_X : \mathbb{R} \to \mathbb{R}$ de la siguiente manera:

$$F_X(x) = P[X \leq x]$$

El problema que se plantea ahora consiste en lo siguiente:

Partiendo de la función de distribución F_X, determinar para qué subconjuntos B de $\mathbb{R}$ es posible encontrar la probabilidad de que X tome valores en el conjunto B, lo cual se representa por $P[X \in B]$.

Vamos a ver que la familia de subconjuntos de $\mathbb{R}$ para los cuales esto es posible, si bien no siempre incluye a todos los subconjuntos de $\mathbb{R}$, es bastante grande y forma una $\sigma-$álgebra.

Podemos pensar este problema de la siguiente forma:

A partir de F_X se puede encontrar fácilmente la probabilidad de que X tome valores en cualquier intervalo $(a,b]$, donde a y b son dos números reales tales que $a<b$; a saber:

$$P\left[a<X\leq b\right]=P\left[X\in(a,b]\right]=F_X\left(b\right)-F_X\left(a\right)$$

Esta relación podemos leerla de la siguiente manera:

El intervalo $(a,b]$ tiene una medida dada por su longitud; es decir, $b-a$; pero en este caso lo que tenemos es $F_X\left(b\right)-F_X\left(a\right)$; así que podemos pensar en esta diferencia como una nueva manera de medir el intervalo $(a,b]$; es decir, la medida de ese intervalo es la probabilidad de que X tome como valor un punto dentro de él. Así que, de manera general, dado un subconjunto B de $\mathbb{R}$, le vamos a asignar como medida la probabilidad de que X tome como valor un elemento de ese conjunto; de manera más precisa, partiendo de la función de distribución F_X, queremos determinar para qué subconjuntos B de $\mathbb{R}$ es posible encontrar $P\left[X\in B\right]$; esta probabilidad será la medida que le asignemos al conjunto B.

Para determinar esos subconjuntos vamos a utilizar las propiedades que tiene cualquier función de distribución F_X, las cuales son las siguientes:

$i)$ F_X es una función no decreciente.

$ii)$ F_X es continua por la derecha.

$iii)$ $\lim_{x\to\infty}F_X\left(x\right)=1$.

$iv)$ $\lim_{x\to-\infty}F_X\left(x\right)=0$.

El problema que queremos resolver se puede plantear de la siguiente manera:

Definamos:

$$F_X(\infty)=\lim_{x\to\infty}F_X(x)=1$$

$$F_X(-\infty)=\lim_{x\to-\infty}F(x)=0$$

Si $a,b\in\mathbb{R}\cup\{-\infty,\infty\}$ y $a<b$, definamos $(a,b|$ de la siguiente manera:

$$(a,b|=\begin{cases}(a,b] & \text{si } b\in\mathbb{R}\\ (a,b) & \text{si } b=\infty\end{cases}$$

Sea $\mathcal{I}$ la familia de los intervalos de este tipo, agregando al vacío como parte de la familia.

Definamos $\mu_X(\emptyset) = 0$ y, para cada intervalo $I = (a, b| \in \mathcal{I}$, $\mu_X(I) = F_X(b) - F_X(a)$.

μ_X es una función definida sobre los intervalos que pertenecen a $\mathcal{I}$.

Se trata ahora de extender μ_X de tal manera que $\mu_X(B)$ quede definida para una familia de subconjuntos B de $\mathbb{R}$ que formen una $\sigma-$álgebra $\Im_X$ tan grande como sea posible y que μ_X sea una medida; es decir, que tenga las siguientes propiedades:

$i)$ $\mu_X(B) \geq 0$ para cualquier $B \in \Im_X$.

$ii)$ Si $(B_n)_{n\in\mathbb{N}}$ es una sucesión de elementos de $\Im_X$ tales que cualquier pareja de ellos está formada por conjuntos ajenos, entonces:

$$\mu_X(\cup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu_X(B_n)$$

El resultado central que nos va a permitir realizar la extensión de μ_X es el siguiente:

Sea $(\mathbf{a}, \mathbf{b}| \in I$ y $(\mathbf{a_1}, \mathbf{b_1}|, (\mathbf{a_2}, \mathbf{b_2}|, \ldots$ una colección infinita de intervalos en I tales que $(\mathbf{a}, \mathbf{b}| \subset \bigcup_{\mathbf{k=1}}^{\infty} (\mathbf{a_k}, \mathbf{b_k}|$. Entonces:

$$\mathbf{F_X(b) - F_X(a) \leq \sum_{k=1}^{\infty} [F_X(b_k) - F_X(a_k)]}$$

Este resultado puede parecer obvio ya que F_X es una función no decreciente, de manera que, si dos intervalos $(a_j, b_j|$ y $(a_k, b_k|$ tienen una intersección no vacía $(a_i, b_i|$, entonces, al sumar $F_X(b_j) - F_X(a_j)$ con $F_X(b_k) - F_X(a_k)$, estaríamos sumando dos veces la diferencia $F_X(b_i) - F_X(a_i)$. Sin embargo, por tratarse de una suma infinita (una serie), en general no hay manera de ordenar los intervalos $(a_k, b_k|$ de manera que, una vez ordenados, podamos hacer ver que al realizar la suma infinita podría haber términos que se sumen varias veces, dando como resultado que la suma $\sum_{k=1}^{\infty} [F_X(b_k) - F_X(a_k)]$ sea mayor que $F_X(b) - F_X(a)$.

Esta dificultad la resolvió Émile Borel, en el año 1895, demostrando un resultado que ahora, en su forma general, se conoce como teorema de Heine Borel, el cual nos permite asegurar lo siguiente:

Si $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$ es un intervalo cerrado y acotado y $((\boldsymbol{\alpha}_\mathbf{n}, \boldsymbol{\beta}_\mathbf{n}))_{\mathbf{n}\in\mathbb{N}}$ es una sucesión de intervalos abiertos tales que $[\boldsymbol{\alpha}, \boldsymbol{\beta}] \subset \cup_{\mathbf{n=1}}^{\infty} (\boldsymbol{\alpha}_\mathbf{n}, \boldsymbol{\beta}_\mathbf{n})$, entonces existe un número finito de esos intervalos abiertos cuya unión cubre al intervalo $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$.

Demostración del resultado central que permite realizar la extensión de μ_X

Lema 1. *Sea* $I = (a, b| \in \mathcal{I}$ *y* $I^{(1)} = \left(a^{(1)}, b^{(1)}\right|, \ldots, I^{(m)} = \left(a^{(m)}, b^{(m)}\right|$ *un conjunto finito de intervalos en* $\mathcal{I}$ *tales que* $I \subset \bigcup_{j=1}^{m} I^{(j)}$*, entonces:*

$$F_X(b) - F_X(a) \leq \sum_{j=1}^{m} \left[F_X(b^{(j)}) - F_X(a^{(j)})\right]$$

Demostración

Los puntos $a, b, a^{(1)}, b^{(1)}, \ldots, a^{(m)}, b^{(m)}$ constituyen una partición de un intervalo $(c, d|$ que contiene al intervalo $(a, b|$.

Esta partición parte cada intervalo $I^{(j)}$, con $j \in \{1, 2, \ldots, m\}$, en subintervalos ajenos por parejas, $\left(c_1^{(j)}, d_1^{(j)}\right|, \left(c_2^{(j)}, d_2^{(j)}\right|, \ldots, \left(c_{n_j}^{(j)}, d_{n_j}^{(j)}\right|$. Así que:

$$F_X(b^{(j)}) - F_X(a^{(j)}) = \sum_{k=1}^{n_j} \left[F_X(d_k^{(j)}) - F_X(c_k^{(j)})\right]$$

La partición definida antes también parte el intervalo I en subintervalos ajenos por parejas, $(c_1, d_1| , (c_2, d_2| , \ldots , (c_n, d_n|$. Así que:

$$F_X(b) - F_X(a) = \sum_{k=1}^{n} \left[F_X(d_k) - F_X(c_k)\right]$$

Por otra parte, como $I \subset \bigcup_{j=1}^{m} I^{(j)}$, cada intervalo $(c_k, d_k|$, con $k \in \{1, 2, \ldots, n\}$, coincide con un intervalo $\left(c_{k'}^{(j)}, d_{k'}^{(j)}\right| \; I_{k'}^{(j)}$ para alguna $j \in \{1, 2, \ldots, m\}$ y alguna $k' \in \{1, 2, \ldots, n_j\}$, por lo tanto:

$F_X(b) - F_X(a) = \sum_{k=1}^{n} \left[F_X(d_k) - F_X(c_k)\right]$

$\leq \sum_{j=1}^{m} \sum_{k=1}^{n_j} \left[F_X(d_k^{(j)}) - F_X(c_k^{(j)})\right] = \sum_{j=1}^{m} \left[F_X(b^{(j)}) - F_X(a^{(j)})\right]$

■

Teorema 1. *Sea* $(a, b| \in \mathcal{I}$ *y* $(a_1, b_1| , (a_2, b_2| , \ldots$ *una colección infinita de intervalos en* $\mathcal{I}$ *tales que* $(a, b| \subset \bigcup_{k=1}^{\infty} (a_k, b_k|$*. Entonces:*

$$F_X(b) - F_X(a) \leq \sum_{k=1}^{\infty} \left[F_X(b_k) - F_X(a_k)\right]$$

Demostración

Tomemos $\varepsilon > 0$ y $\delta > 0$, arbitrarios.

Como F_X es continua por la derecha, para cada cada $k \in \mathbb{N}$, existe $\delta_k > 0$ tal que:

$$F_X\left(d_k\right) - F_X\left(b_k\right) < \tfrac{\varepsilon}{2^k},$$

donde:

$$d_k = \begin{cases} b_k + \delta_k & \text{si } b_k \in \mathbb{R} \\ b_k & \text{si } b_k = \infty \end{cases}$$

Definamos:

$$c_\delta = \begin{cases} a + \delta & \text{si } a \in \mathbb{R} \\ -\frac{1}{\delta} & \text{si } a = -\infty \end{cases}$$

$$d_\delta = \begin{cases} b & \text{si } b \in \mathbb{R} \\ \frac{1}{\delta} & \text{si } b = \infty \end{cases}$$

Entonces:

$$\text{lím}_{\delta \to 0} \left[F_X(d_\delta) - F_X(c_\delta) \right] = F_X(b) - F_X(a)$$

Además:

$$(c_\delta, d_\delta| \subset [c_\delta, d_\delta] \subset (a, b| \subset \bigcup_{k=1}^{\infty} (a_k, b_k| \subset \bigcup_{k=1}^{\infty} (a_k, d_k)$$

Así que, por el teorema de Heine-Borel, existe una colección finita, $(a_{k_1}, d_{k_1}), \ldots, (a_{k_m}, d_{k_m})$, tal que:

$$[c_\delta, d_\delta] \subset \bigcup_{j=1}^{m} \left(a_{k_j}, d_{k_j} \right)$$

Por lo tanto:

$$(c_\delta, d_\delta| \subset [c_\delta, d_\delta] \subset \bigcup_{j=1}^{m} \left(a_{k_j}, d_{k_j} \right) \subset \bigcup_{j=1}^{m} \left(a_{k_j}, d_{k_j} \right|$$

Así que:

$F_X(d_\delta) - F_X(c_\delta) \leq \sum_{j=1}^{m} \left[F_X\left(d_{k_j}\right) - F_X\left(a_{k_j}\right) \right] \leq \sum_{k=1}^{\infty} \left[F_X(d_k) - F_X(a_k) \right]$

$\leq \sum_{k=1}^{\infty} \left[F_X(b_k) - F_X(a_k) \right] + \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \sum_{k=1}^{\infty} \left[F_X(b_k) - F_X(a_k) \right] + \varepsilon$

Y, como $\varepsilon > 0$ es arbitraria:

$$F_X(d_\delta) - F_X(c_\delta) \leq \sum_{k=1}^{\infty} \left[F_X(b_k) - F_X(a_k) \right]$$

Finalmente, tomando límites cuando $\delta \to 0$, se obtiene:

$$F_X(b) - F_X(a) \leq \sum_{k=1}^{\infty} \left[F_X(b_k) - F_X(a_k) \right]$$

$\blacksquare$

Con lo anterior ya contamos con lo necesario para poder extender μ_X a una $\sigma-$álgebra de subconjuntos de $\mathbb{R}$.

Los pasos a seguir para realizar esta extensión son los siguientes:

1. Vamos a extender μ_X a una familia $\mathcal{A}$ de subconjuntos de $\mathbb{R}$ qué aún no forman una $\sigma-$álgebra, pero sí un álgebra de subconjuntos de $\mathbb{R}$, la cual se define como una familia de conjuntos que tiene las siguientes propiedades:

$i)$ $\mathbb{R} \in \mathcal{A}$.

$ii)$ Si $A \in \mathcal{A}$, entonces $A^c \in \mathcal{A}$.

$iii)$ Si $A_1, A_2, \ldots, A_n$ es una colección finita de elementos de $\mathcal{A}$, entonces $\cup_{k=1}^n A_k \in \mathcal{A}$.

2. Utilizando el resultado central que demostramos antes, vamos a demostrar que μ_X tiene la siguiente propiedad:

Si $A_1, A_2, \ldots$ es una colección infinita numerable de elementos de $\mathcal{A}$, ajenos por parejas, no vacíos y tales que $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$, entonces:

$$\mu_X \left(\bigcup_{n=1}^\infty A_i\right) = \sum_{n=1}^\infty \mu_X \left(A_n\right)$$

3. Como corolario, probaremos que, dada cualquier colección infinita $A_1, A_2, \ldots$ de elementos de $\mathcal{A}$ tales que $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$, entonces:

$$\mu_X \left(\bigcup_{n=1}^\infty A_n\right) \leq \sum_{n=1}^\infty \mu_X(A_n)$$

4. Para extender μ_X a una $\sigma-$álgebra de subconjuntos de $\mathbb{R}$, vamos a definir primero, para cualquier subconjunto A de $\mathbb{R}$, lo que se llama una medida exterior; es decir, nos vamos a acercar al conjunto A con uniones infinitas numerables de elementos de $\mathcal{A}$ que cubran A..

5. Demostraremos algunas propiedades de la medida exterior previamente definida.

6. Caracterizaremos a los subconjuntos de $\mathbb{R}$ para los cuales se puede definir su medida, dándoles el nombre de conjuntos medibles y asignándoles como medida μ_X su medida exterior.

7. Mostraremos que cualquier elemento de $\mathcal{A}$ es medible.

8. Demostraremos que la familia $\Im_X$ de los conjuntos medibles forma un álgebra de subconjuntos de $\mathbb{R}$.

9. Probaremos que μ_X, definida sobre $\Im_X$, es finitamente aditiva; es decir si $E_1, E_2, \ldots, E_n$ son conjuntos medibles, ajenos por parejas, entonces:

$$\mu_X \left(U_{k=1}^n E_k \right) = \sum_{k=1}^n \mu_X \left(E_k \right)$$

Para esto, bastará con probar que si E_1 y E_2 son conjuntos medibles ajenos, entonces $\mu_X \left(E_1 \cup E_2 \right) = \mu_X \left(E_1 \right) + \mu_X \left(E_2 \right)$, ya que, habiendo demostrado esto, se puede probar que se tiene una propiedad similar para 3 conjuntos medibles ajenos por parejas; después para 4 conjuntos, y así sucesivamente.

10. Demostraremos que la familia $\Im_X$ de los conjuntos medibles forma una $\sigma-$álgebra de subconjuntos de $\mathbb{R}$.

11. Probaremos que μ_X, definida sobre $\Im_X$, es $\sigma-$aditiva, con lo cual habremos demostrado que es una medida definida sobre una $\sigma-$álgebra de subconjuntos de $\mathbb{R}$.

Extensión de μ_X a un álgebra de subconjuntos de $\mathbb{R}$

Sea $\mathcal{A}$ la familia formada por los conjuntos de la forma $\bigcup_{j=1}^n I_j$ donde $n \in \mathbb{N}$ y $I_1, \ldots, I_n$ son intervalos en $\mathcal{I}$, ajenos por parejas.

$\mathcal{A}$ es un álgebra de subconjuntos de $\mathbb{R}$.

Para cada $A = \bigcup_{j=1}^n I_j \in \mathcal{A}$, definamos $\mu_X(A) = \sum_{j=1}^n \mu_X(I_j)$.

Mostremos que μ_X está bien definida; es decir, que si A se puede expresar de diferentes maneras como una unión finita de intervalos en $\mathcal{I}$, ajenos por parejas, se obtiene el mismo valor para $\mu_X(A)$.

Lema 2. *Sea $I = (a, b| \in \mathcal{I}$ y $\left(a^{(1)}, b^{(1)}\right|, \left(a^{(2)}, b^{(2)}\right|, \ldots, \left(a^{(m)}, b^{(m)}\right|$, una colección finita de intervalos en $\mathcal{I}$, ajenos por parejas, tales que $I = \bigcup_{j=1}^m \left(a^{(j)}, b^{(j)}\right|$, entonces:*

$$\mu_X \left(I \right) = \sum_{j=1}^m \mu_X \left(\left(a^{(j)}, b^{(j)}\right| \right)$$

Demostración

Como los intervalos $\left(a^{(1)}, b^{(1)}\right|, \left(a^{(2)}, b^{(2)}\right|, \ldots, \left(a^{(m)}, b^{(m)}\right|$ son ajenos por parejas y su unión es $(a, b|$, podemos ordenarlos para obtener una colección de intervalos ajenos por parejas, $\left(x^{(1)}, y^{(1)}\right|, \left(x^{(2)}, y^{(2)}\right|, \ldots, \left(x^{(m)}, y^{(m)}\right|$, de tal forma que:

$$a = x^{(1)} < y^{(1)} = x^{(2)} < y^{(2)} = x^{(2)} < \cdots = x^{(m)} < y^{(m)} = b$$

Así que:

$\sum_{j=1}^{m} \mu_X \left(\left(a^{(j)}, b^{(j)}\right|\right) = \sum_{j=1}^{m} \mu_X \left(\left(y^{(j)}, z^{(j)}\right|\right) = \sum_{j=1}^{m} \left[F_X \left(y^{(j)}\right) - F_X \left(x^{(j)}\right)\right]$

$= F_X (b) - F_X (a) = \mu_X (I)$

■

Lema 3. *Sean $I_1, \ldots, I_k$ y $I^{(1)}, \ldots, I^{(m)}$ dos colecciones finitas de intervalos en $\mathcal{I}$ tales que $I_1, \ldots, I_k$ son ajenos por parejas, $I^{(1)}, \ldots, I^{(m)}$ son ajenos por parejas y $\bigcup_{i=1}^{k} I_i = \bigcup_{j=1}^{m} I^{(j)}$, entonces:*

$$\sum_{i=1}^{k} \mu_X (I_i) = \sum_{j=1}^{m} \mu_X \left(I^{(j)}\right)$$

Demostración

Para cada $i \in \{1, \ldots, k\}$ y $j \in \{1, \ldots, m\}$, definamos $I_i^{(j)} = I_i \cap I^{(j)}$. Entonces, como $\bigcup_{i=1}^{k} I_i = \bigcup_{j=1}^{m} I^{(j)}$, se tiene $I_i = \bigcup_{j=1}^{m} I_i^{(j)}$ y $I^{(j)} = \bigcup_{i=1}^{k} I_i^{(j)}$, así que:

$$\mu_X (I_i) = \sum_{j=1}^{m} \mu_X \left(I_i^{(j)}\right)$$

$$\mu_X \left(I^{(j)}\right) = \sum_{i=1}^{k} \mu_X \left(I_i^{(j)}\right)$$

Por lo tanto:

$\sum_{i=1}^{k} \mu_X (I_i) = \sum_{i=1}^{k} \sum_{j=1}^{m} \mu_X \left(I_i^{(j)}\right)$

$= \sum_{j=1}^{m} \sum_{i=1}^{k} \mu_X \left(I_i^{(j)}\right) = \sum_{j=1}^{m} \mu_X \left(I^{(j)}\right)$

■

Obviamente la función $\mu_X : \mathcal{A} \mapsto \mathbb{R}$ es no negativa y finitamente aditiva.

Teorema 2. *Sea $A_1, A_2, \ldots$ una colección infinita numerable de elementos de $\mathcal{A}$, ajenos por parejas, no vacíos y tales que $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, entonces:*

$$\mu_X \left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu_X (A_i)$$

Demostración

Sea $A = \bigcup_{i=1}^{\infty} A_i$.

Como para cada $i \in \mathbb{N}$, $A_i \in \mathcal{A}$, A_i es una unión finita de intervalos en $\mathcal{I}$ ajenos por parejas. Además, como $A \in \mathcal{A}$, A también es una unión finita de intervalos en $\mathcal{I}$ ajenos por parejas.

Sean $A = \bigcup_{j=1}^{m} I^{(j)}$ y, para cada $i \in \mathbb{N}$, $A_i = \bigcup_{k=1}^{m_i} I_{(i,k)}$. Entonces:

$$\bigcup_{j=1}^{m} I^{(j)} = A = \bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} \bigcup_{k=1}^{m_i} I_{(i,k)}$$

Tenemos dos colecciones de intervalos, por un lado la familia $\left\{I_{(i,k)} : i \in \mathbb{N}, k \in \{1, \ldots, m_i\}\right\}$ y por el otro la familia $\left\{I^{(j)} : j \in \{1, \ldots, m\}\right\}$. Tanto los intervalos de la primera familia como los de la segunda son ajenos por parejas y A es igual tanto a la unión de los intervalos de la primera familia como de la segunda. Por otra parte, una pareja de intervalos, uno de la primera familia y otro de la segunda, podrían no ser ajenos.

La idea ahora es partir cada intervalo $I^{(j)}$ en intervalos ajenos por parejas, utilizando los intervalos $I_{(i,k)}$. Para esto, definamos, para cada $j \in \{1, \ldots, m\}$, $i \in \mathbb{N}$ y $k \in \{1, \ldots, m_i\}$:

$$I_{(i,k)}^{(j)} = I_{(i,k)} \cap I^{(j)}$$

Entonces, los intervalos de la familia $\left\{I_{(i,k)}^{(j)} : j \in \{1, \ldots, m\}, i \in \mathbb{N}, k \in \{1, \ldots, m_i\}\right\}$ son ajenos por parejas y:

$I_{(i,k)} = \bigcup_{j=1}^{m} I_{(i,k)}^{(j)}$ para cualesquiera $i \in \mathbb{N}$ y $k \in \{1, \ldots, m_i\}$.

$I^{(j)} = \bigcup_{i=1}^{\infty} \bigcup_{k=1}^{m_i} I_{(i,k)}^{(j)}$ para cualquier $j \in \{1, \ldots, m\}$.

Además, por el teorema 1, se tiene:

$\mu_X\left(I^{(j)}\right) \leq \sum_{i=1}^{\infty} \sum_{k=1}^{m_i} \mu_X\left(I_{(i,k)}^{(j)}\right)$ para cualquier $j \in \{1, \ldots, m\}$.

Así que:

$\mu_X(A) = \sum_{j=1}^{m} \mu_X\left(I^{(j)}\right) \leq \sum_{j=1}^{m} \sum_{i=1}^{\infty} \sum_{k=1}^{m_i} \mu_X\left(I_{(i,k)}^{(j)}\right)$

$= \sum_{i=1}^{\infty} \sum_{k=1}^{m_i} \sum_{j=1}^{m} \mu_X\left(I_{(i,k)}^{(j)}\right) = \sum_{i=1}^{\infty} \sum_{k=1}^{m_i} \mu_X\left(I_{(i,k)}\right)$

$= \sum_{i=1}^{\infty} \mu_X(A_i)$

Además, como μ_X es finitamente aditiva y $A \supset \bigcup_{i=1}^{n} A_i$ para cualquier $n \in \mathbb{N}$, se tiene $\mu_X(A) \geq \sum_{i=1}^{n} \mu_X(A_i)$ para cualquier $n \in \mathbb{N}$, así que:

$$\mu_X(A) \geq \sum_{i=1}^{\infty} \mu_X(A_i)$$

Por lo tanto, $\mu_X(A) = \sum_{i=1}^{\infty} \mu_X(A_i)$.

$\blacksquare$

Corolario 1. *Sea* $A_1, A_2, \ldots$ *una colección infinita numerable de elementos de* $\mathcal{A}$ *tales que* $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, *entonces:*

$$\mu_X\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu_X(A_n)$$

Demostración

Para cada $n \in \mathbb{N}$, sea $B_n = A_n - \bigcup_{j=1}^{n-1} A_j$, entonces los conjuntos $B_1, B_2, \ldots$ son ajenos por parejas y $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$, así que:

$$\mu_X\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu_X\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu_X(B_n) \leq \sum_{n=1}^{\infty} \mu_X(A_n)$$

■

Extensión de μ_X a una $\sigma-$álgebra de subconjuntos de $\mathbb{R}$

Definición 1. *Diremos que una colección finita o infinita numerable* $A_1, A_2, \ldots$ *de elementos de* $\mathcal{A}$ *es una cubierta de un conjunto* $A \subset \mathbb{R}$ *si* $A \subset \bigcup_n A_n$.

Definición 2. *Se define la medida exterior,* $\mu_e(A)$, *de un subconjunto* A *de* $\mathbb{R}$, *mediante la relación:*

$$\mu_e(A) = \text{ínf}\left\{\sum_j \mu_X(A_j) : A_1, A_2, \ldots \text{ es cubierta de } A\right\}$$

Definición 3. *Diremos que un conjunto* E *es medible si* $\mu_e(A) = \mu_e(A \cap E) + \mu_e(A \cap E^c)$ *para cualquier conjunto* A. *Además, en este caso, se define* $\mu_X(E)$, *como la medida exterior de* E.

Denotaremos por $\Im_X$ a la familia de conjuntos medibles.

NOTA AL MARGEN

Quienes iniciaron el desarrollo de la teoría de la medida, como se le conoce actualmente, fueron los matemáticos franceses Émile Borel y Henri Lebesgue a principios del siglo XX. Previamente se contaba son la teoría del contenido, del también matemático francés Camille Jordan, la cual se encuentra estrechamente vinculada con la teoría de integración de Riemann. El objetivo de Lebesgue al inventar un método para construir medidas no fue propiamente desarrollar una teoría de la medida; lo que buscaba era una nueva definición de la integral de una función la cual ampliara la familia de funciones integrables y tuviera mejores propiedades que la integral de Riemann. Es por eso que la teoría de la medida se encuentra estrechamente vinculada con esa nueva definición de la integral, que es conocida actualmente como integral de Lebesgue.

En el desarrollo de su trabajo, Lebesgue se enfrentó al problema de definir el concepto de longitud de cualquier subconjunto de los números reales y el método que siguió para resolverlo fue básicamente el mismo que estamos utilizando ahora para generar una medida a partir de una función de distribución. Para definir los conjuntos medibles, los cuales serían aquellos a los cuales se les pueda asignar una longitud, siguió el método que había desarrollado Jordan; a saber, dado cualquier subconjunto acotado A de $\mathbb{R}$, definió la medida exterior $m_e(A)$ y la medida interior $m_i(A)$ del conjunto; la primera acercándose por afuera del conjunto (mediante intervalos) y la segunda, acercándose por adentro. Definió entonces un conjunto E como medible si su medida interior y su medida exterior coinciden. Este método suena bastante natural ya que, al coincidir la medida por dentro del conjunto y su medida por fuera, puede pensarse que, la medida que resulta es exactamente la que corresponde a todo el conjunto E, sin que sobre o falte algo.

En el año 1914, el matemático alemán Constantin Carathéodory inventó otro método para definir a los conjuntos medibles, el cual, además de poderse aplicar de manera general a otros problemas donde se quisiera definir una medida, permite un manejo más simple de la medibilidad y, si se aplica al caso que trató Lebesgue, resulta ser equivalente a la definición que él dio.

La definición de Carathéodory fue la que prevaleció y es la que estamos utilizando en el problema que nos hemos planteado.

Una manera informal de visualizar los conjuntos medibles, de acuerdo con la definición de Carathéodory, es básicamente lo que dijimos antes con relación a la definición de Lebesgue; a saber, consiste en verlos como aquellos conjuntos E cuya medida exterior mide exactamente lo que está dentro del conjunto; es decir, no mide algo más grande que lo que está dentro de él, ni algo más pequeño; esto es lo que estaría diciendo la definición $\mu_X(E) = \mu_e(E)$ para los conjuntos medibles, ya que la propiedad $\mu_e(A) = \mu_e(A \cap E) + \mu_e(A \cap E^c)$, válida para cualquier subconjunto A de $\mathbb{R}$, nos dice que al calcular $\mu_e(A \cap E)$ y $\mu_e(A \cap E^c)$, E y E^c aportan exactamente lo que está dentro de cada uno de esos conjuntos, de manera que al efectuar la suma $\mu_e(A \cap E) + \mu_e(A \cap E^c)$ se obtiene exactamente lo que corresponde a la medida exterior de A.

A continuación veremos algunas de las propiedades de la medida exterior que hemos definido.

Se puede ver inmediatamente que Si A y B son dos subconjuntos de $\mathbb{R}$ tales que $A \subset B$ entonces $\mu_e(A) \leq \mu_e(B)$.

Proposición 1. *Si $A \in \mathcal{A}$ entonces $\mu_e(A) = \mu_X(A)$.*

Demostración

Sea $A \in \mathcal{A}$ y $A_1, A_2, \ldots$ una cubierta de A, entonces $A_n \cap A \in \mathcal{A}$ para cualquier elemento A_n de la cubierta y $\bigcup_n (A_n \cap A) = A \in \mathcal{A}$; así que, por el corolario 1:

$$\mu_X(A) = \mu_X(\bigcup_n (A_n \cap A)) \leq \sum_n \mu_X(A_n \cap A) \leq \sum_n \mu_X(A_n)$$

Por lo tanto, como esto ocurre para cualquier cubierta de A, $\mu_X(A) \leq \mu_e(A)$.

Por otra parte, como A es una cubierta de él mismo, se tiene $\mu_e(A) \leq \mu_X(A)$.

Así que, $\mu_e(A) = \mu_X(A)$.

■

Corolario 2. *$\mu_e(A) \leq 1$ para cualquier subconjunto A de $\mathbb{R}$.*

Demostración

Como $\mathbb{R} \in \mathcal{A}$, se tiene:

$$\mu_e(\mathbb{R}) = \mu_X(\mathbb{R}) = \mu_X((-\infty, \infty)) = F_X(\infty) - F_X(-\infty) = 1$$

Así que, para cualquier conjunto $A \subset \mathbb{R}$, se tiene:

$$\mu_e(A) \leq \mu_e(\mathbb{R}) = 1$$

■

Proposición 2. *Si $(A_n)_{n\in\mathbb{N}}$ es una sucesión de subconjuntos de $\mathbb{R}$, entonces:*

$$\mu_e\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu_e(A_n)$$

Demostración

Dada $\varepsilon > 0$, para cada conjunto A_n, sea $A_n^{(1)}, A_n^{(2)}, \ldots$ una cubierta de A_n tal que $\sum_m \mu_X(A_n^{(m)}) < \mu_e(A_n) + \frac{\varepsilon}{2^n}$.

La familia de conjuntos $A_n^{(m)}$ forman una cubierta de $\bigcup_n A_n$, así que:

$$\mu_e\left(\bigcup_n A_n\right) \leq \sum_n \sum_m \mu_X(A_n^{(m)}) \leq \sum_n \left[\mu_e(A_n) + \frac{\varepsilon}{2^n}\right] \leq \sum_n \mu_e(A_n) + \varepsilon$$

Es decir, $\mu_e\left(\bigcup_n A_n\right) \leq \sum_n \mu_e(A_n) + \varepsilon$ para cualquier $\varepsilon > 0$. Por lo tanto:

$$\mu_e\left(\bigcup_n A_n\right) \leq \sum_n \mu_e(A_n)$$

■

A la propiedad de la medida exterior, demostrada en la proposición anterior, se le llama $\sigma-$**subaditividad**.

Obsérvese que, por esta propiedad, se tiene:

$$\mu_e(A) \leq \mu_e(A \cap E) + \mu_e(A \cap E^c)$$

para cualquier par de conjuntos E y A, de manera que para demostrar la medibilidad de un conjunto E únicamente es necesario probar la otra desigualdad.

Proposición 3. *Todo elemento de $\mathcal{A}$ es medible.*

Demostración

Sean $E \in \mathcal{A}$, A cualquier conjunto y $A_1, A_2, \ldots$ una cubierta de A, entonces, para cada A_n, los conjuntos $A_n \cap E$ y $A_n \cap E^c$ pertenecen a $\mathcal{A}$ y se tiene:

$\mu_e(A \cap E) \leq \mu_e\left(\left(\bigcup_n A_n\right) \cap E\right) = \mu_e\left(\bigcup_n (A_n \cap E)\right)$

$\leq \sum_n \mu_e(A_n \cap E) = \sum_n \mu_X(A_n \cap E)$

$\mu_e(A \cap E^c) \leq \mu_e\left(\left(\bigcup_n A_n\right) \cap E^c\right) = \mu_e\left(\bigcup_n (A_n \cap E^c)\right)$

$\leq \sum_n \mu_e(A_n \cap E^c) = \sum_n \mu_X(A_n \cap E^c)$

Así que:

$$\mu_e(A \cap E) + \mu_e(A \cap E^c) \leq \sum_n \mu_X(A_n \cap E) + \sum_n \mu_X(A_n \cap E^c) = \sum_n \mu_X(A_n)$$

Finalmente, como lo anterior es válido para cualquier cubierta de A, se puede concluir que:

$$\mu_e(A \cap E) + \mu_e(A \cap E^c) \leq \mu_e(A)$$

■

Proposición 4. *La familia de conjuntos medibles forma un álgebra de subconjuntos de $\mathbb{R}$.*

Demostración

Que el conjunto $\mathbb{R}$ es medible, así como que el complemento de un conjunto medible es medible, son resultados obvios.

Sean E_1 y E_2 dos conjuntos medibles y A cualquier conjunto. Se tiene entonces:

$\mu_e(A \cap (E_1 \cup E_2)) + \mu_e(A \cap (E_1 \cup E_2)^c)$

$= \mu_e((A \cap E_1) \cup (A \cap E_1^c \cap E_2)) + \mu_e(A \cap E_1^c \cap E_2^c)$

$\leq \mu_e(A \cap E_1) + \mu_e(A \cap E_1^c \cap E_2) + \mu_e(A \cap E_1^c \cap E_2^c)$

$= \mu_e(A \cap E_1) + \mu_e(A \cap E_1^c) = \mu_e(A)$

Así que, $E_1 \cup E_2$ es medible.

■

Proposición 5. *La función que asigna a cada conjunto medible E su medida, $\mu_X(E)$, es una función finitamente aditiva.*

Demostración

Sean E_1 y E_2 dos conjuntos medibles ajenos, entonces, como $E_1 \cup E_2$ es medible, se tiene:

$\mu(E_1 \cup E_2)) = \mu((E_1 \cup E_2)) \cap E_1) + \mu((E_1 \cup E_2) \cap E_1^c)$

$= \mu(E_1) + \mu(E_2)$

■

Proposición 6. *La familia de conjuntos medibles forma una σ-álgebra de subconjuntos de $\mathbb{R}$.*

Demostración

Sea $E_1, E_2, \ldots$ una colección infinita numerable de conjuntos medibles ajenos por parejas y A cualquier subconjunto de $\mathbb{R}$.

Demostremos que $\mu_e\left(A \cap (\bigcup_{j=1}^{n} E_j)\right) = \sum_{j=1}^{n} \mu_e(A \cap E_j)$ para cualquier $n \in \mathbb{N}$.

Para $n = 1$ la igualdad es obvia.

Supongamos ahora que la igualdad es válida para $n = k$, entonces, como E_{k+1} es medible, se tiene:

$\mu_e\left(A \cap (\bigcup_{j=1}^{k+1} E_j)\right) = \mu_e\left(A \cap (\bigcup_{j=1}^{k+1} E_j) \cap E_{k+1}\right) + \mu_e\left(A \cap (\bigcup_{j=1}^{k+1} E_j) \cap E_{k+1}^c\right)$

$= \mu_e(A \cap E_{k+1}) + \mu_e\left(A \cap (\bigcup_{j=1}^{k} E_j)\right) = \mu_e(A \cap E_{k+1}) + \sum_{j=1}^{k} \mu_e(A \cap E_j)$

$= \sum_{j=1}^{k+1} \mu_e(A \cap E_j)$

Por lo tanto, la igualdad es válida para $n = k + 1$, así que, por el principio de inducción, lo es para cualquier $n \in \mathbb{N}$.

Ahora bien, como la familia de conjuntos medibles forma un álgebra de subconjuntos de $\mathbb{R}$, para cada $n \in \mathbb{N}$ el conjunto $\bigcup_{j=1}^{n} E_j$ es medible, así que:

$\mu_e(A) = \mu_e\left(A \cap (\bigcup_{j=1}^{n} E_j)\right) + \mu_e\left(A \cap (\bigcup_{j=1}^{n} E_j)^c\right)$

$= \sum_{j=1}^{n} \mu_e(A \cap E_j) + \mu_e\left(A \cap (\bigcup_{j=1}^{n} E_j)^c\right)$

$\geq \sum_{j=1}^{n} \mu_e(A \cap E_j) + \mu_e\left(A \cap (\bigcup_{j=1}^{\infty} E_j)^c\right)$

Tomando límite cuando $n \to \infty$ y utilizando la σ-subaditividad de la medida exterior,se obtiene entonces:

$\mu_e(A) \geq \sum_{j=1}^{\infty} \mu_e(A \cap E_j) + \mu_e\left(A \cap (\bigcup_{j=1}^{\infty} E_j)^c\right)$

$\geq \mu_e\left(A \cap (\bigcup_{j=1}^{\infty} E_j)\right) + \mu_e\left(A \cap (\bigcup_{j=1}^{\infty} E_j)^c\right)$

Por lo tanto, $\bigcup_{j=1}^{\infty} E_j$ es medible.

■

Proposición 7. *La función que asigna a cada conjunto medible E su medida, $\mu_X(E)$, es una función σ-aditiva.*

Demostración

Sea $E_1, E_2, \ldots$ una colección infinita numerable de conjuntos medibles ajenos por parejas. Por la σ-subaditividad de la medida exterior, se tiene:

$$\mu_X(\bigcup_{j=1}^{\infty} E_j) \leq \sum_{j=1}^{\infty} \mu_X(E_j)$$

Por otra parte, por la aditividad finita de la función que asigna a cada conjunto medible su medida, se tiene, para cualquier $n \in \mathbb{N}$:

$$\mu_X(\bigcup_{j=1}^{\infty} E_j) \geq \mu_X(\bigcup_{j=1}^{n} E_j) = \sum_{j=1}^{n} \mu_X(E_j)$$

Así que tomando límite cuando $n \to \infty$, se tiene:

$$\mu_X(\bigcup_{j=1}^{\infty} E_j) \geq \sum_{j=1}^{\infty} \mu_X(E_j)$$

Por lo tanto:

$$\mu_X(\bigcup_{j=1}^{\infty} E_j) = \sum_{j=1}^{\infty} \mu_X(E_j)$$ ■

Los resultados anteriores pueden condensarse en el siguiente teorema:

Teorema 3. *Sean* $(\Omega, \Im, P)$ *un espacio de probabilidad,* $X : \Omega \to \mathbb{R}$ *una variable aleatoria y* $F_X : \mathbb{R} \to \mathbb{R}$ *la función de distribución de* X*. Entonces existe una medida* μ_X*, definida sobre una* $\sigma-$*álgebra* $\Im_X$ *de subconjuntos de* $\mathbb{R}$*, la cual contiene a los intervalos de la forma* $(a, b]$*, donde* $a, b \in \mathbb{R}$ *y* $a < b$*, y tal que, si* $(a, b]$ *es cualquiera de esos intervalos, entonces:*

$$\mu_X((a, b]) = F_X(b) - F_X(a)$$

Obsérvese que μ_X es una medida de probabilidad, ya que $\mu_X(\mathbb{R}) = 1$.

Así como $F_X(b) - F_X(a)$ es igual a la probabilidad de que X tome como valor algún elemento del intervalo $(a, b]$, si B es cualquier elemento de $\Im_X$, $\mu_X(B)$ es igual a la probabilidad de que X tome como valor algún elemento del conjunto B.

Así que podemos concluir que **la función de distribución $\mathbf{F_X}$ de una variable aleatoria $\mathbf{X}$ representa una medida de probabilidad $\boldsymbol{\mu_X}$, definida sobre una $\boldsymbol{\sigma}-$álgebra $\Im_{\mathbf{X}}$ de subconjuntos de $\mathbb{R}$, la cual contiene a los intervalos de la forma $(\mathbf{a}, \mathbf{b}]$, donde $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ y $\mathbf{a} < \mathbf{b}$, y tal que, si $\mathbf{B} \in \Im_{\mathbf{X}}$, entonces:**

$$\mathbf{P}[\mathbf{X} \in \mathbf{B}] = \boldsymbol{\mu}_{\mathbf{X}}(\mathbf{B})$$

Obsérvese que, al ser $\Im_X$ una $\sigma-$álgebra que contiene a los intervalos de la forma $(a, b]$, donde $a, b \in \mathbb{R}$ y $a < b$, se deduce que $\Im_X$ contiene a todos los intervalos, del tipo que sean, ya que, utilizando las operación de unión numerable, intersección numerable y complemento, cualquier intervalo se puede obtener a partir de intervalos de la forma $(a, b]$. Por ejemplo:

$$(a, b) = \cup_{n=1}^{\infty} \left(a, b - \tfrac{1}{n}\right]$$

$$[a, b] = \cap_{n=1}^{\infty} \left(a - \tfrac{1}{n}, b\right]$$